Final Performance Report

Grant Number: PW-253734-17

White Paper:
The Georgian Papers Programme: Transatlantic Access and Discovery Planning Stage

Project Director: Karin Wulf

Grantee Institution: Omohundro Institute of Early American History and Culture

White Paper Submission Date: July 31, 2018

A. Project Activities

The Georgian Papers Programme (GPP) is a partnership between Royal Collection Trust (RCT) and King's College London (King's) and primary U.S. partners the Omohundro Institute of Early American History & Culture (OI) and William & Mary.

Launched in 2015, the GPP is an interdisciplinary project to digitize, conserve, catalog, transcribe, interpret and disseminate items in the Royal Archives and Royal Library relating to the Georgian period, 1714–1837. In conjunction with academics at King's, the OI, and William & Mary, and with other participating U.S. institutions including Mount Vernon and the Library of Congress, the GPP will digitize and disseminate, in simultaneous stages of discovery, access and interpretation, the archive of the Georgian monarchs held in the Royal Archives at Windsor Castle.

The NEH HCRR grant allowed the OI, working with the other primary partners—William & Mary, King's and RCT—to engage in strategic planning over a twelve-month period from May 2017–April 2018 about the best workflow between the academic and technical activities of this significant transatlantic project.

When the GPP launched, the GPP partners were aware that the size of the collection could only be estimated; it is now determined that the collection includes close to 425,000 pages and 65,000 items. The ultimate goal of the GPP is to provide a unique digital resource that is both readily accessible to members of the public and capable of sophisticated manipulation by researchers in any discipline and to offer academic and public programming representing fresh research and interpretation. Over 60,000 pages of the historic documents have been published already with more than 20,000 new pages of material due to be launched in late summer 2018.

The GPP is part of a wider program by the RCT, which has responsibility for the Royal Library and Royal Archives, to open access to its primary source material. Because of the practical limitations of providing access for researchers at Windsor, the GPP's first goal is to digitize the materials documenting the Hanoverian Dynasty, dating from 1714–1837, and offer an online archive and library available to all—academics and the public. By 2021, the GPP will scan all of the material in a new, specially-designed digital lab at Windsor Castle. Selections of archive material are released in scheduled, semi-annual tranches of 30,000-60,000 pages as JPEG files with basic cataloging.

The GPP's second goal, undertaken by teams at King's, the OI and William & Mary, is to use digital technologies to speed the transcription of these original texts and simultaneously produce extensive metadata that will encompass the interests of public audiences as well as academic research projects. One result of the NEH HCRR planning grant is to near completion of transcription of all of the King George III essays. The transcribed essays will be available by late summer 2018; this corpus will serve as an example of transcription modes and an opportunity for potential digital humanities work in the GPP.

The GPP's third goal is to provide all users access to the texts and related metadata through enriched searching, text mining and data visualization. All three goals depend on work carried

out by interdisciplinary teams of librarians, archivists, digital specialists and scholars engaged in interlocking stages of digitization, transcription and interpretation.

Official papers of this period have been available in governmental records at the National Archives at Kew, but the Royal Archives collection provides an entirely different and distinctive resource for the Georgian period. Only about 15 percent of the approximately 425,000 pages of GPP materials have previously been published and little more had been thoroughly examined by scholars. While the vast majority of the collection is from the reign of George III, the papers from the eras of Kings George I, George II, George IV and William IV will also be significant for historical scholarship.

For scholars of early America and the Atlantic world, the Georgian Papers offer a fresh, unparalleled look at the worlds and people who were part of the cultural, economic, military and political machinery of empire. The early exploration by more than 100 scholars, archivists, and librarians since the project's launch in April 2015 has confirmed the significance of the Georgian Papers for understanding political and military aspects of the American Revolution as well as other, wide-ranging aspects of Atlantic, North American and English and European history. Because the materials had not been previously cataloged and are only now beginning to be more thoroughly identified under the auspices of the GPP, it is still not clear beyond the top-level chronological organization by monarchical reign how the collection will relate to manuscripts in the U.K. National Archives or elsewhere. The work of archivists and librarians in conjunction with the first groups of scholars using the materials have begun to illuminate more detail about their contents (details below).

With the support of an NEH HCRR planning grant and funding from GPP partners, teams at King's, William & Mary, and the OI, developed research models for sharing information about the collection content. We then used advance, on-site and remote access to the pre-release materials to create and evaluate a pilot for building technical infrastructure and creating processes for transcription and metadata enrichment to allow for robust digital humanities research and analysis.

Major activities during the NEH HCRR grant period included the following (with additional detail in B. Accomplishments):

- In Fall 2017, GPP fellows gathered on September 4 at Windsor for a symposium to present their research and work in the archives for the benefit of other scholars, librarians, archivists, and technologists working on the GPP. (https://georgianpapers-us.wm.edu/conferences-workshops/workshops/).
- The following day, September 5, 2017, a workshop focused on transcription and platform requirements took place at King's College London with research fellows, all project partners and invited experts in the digital field in attendance.
- Websites in the U.S. (https://gpp.wm.edu/) and U.K (http://georgianpapersprogramme.com/) continue to provide academic interpretation and information in concert with the RCT site to keep researchers informed about developments. Plans for the coming year include merging the U.S. and U.K sites into one web destination.
- On December 1, 2017, a GPP symposium was held at the Library of Congress's Kluge

Center to provide scholars and the general public with an opportunity to hear from GPP partners and fellows who have worked in the archives and learn more about the opportunity provided by this rich resource.

- Throughout the grant period, the GPP partners met regularly, including the quarterly meetings of the Programme Steering Committee (PSC) as well as the subcommittees of the PSC: Academic Steering, Operations, and Communications and Development.
- During fellows' coffee mornings at King's, all GPP fellows (including the OI-GPP fellows) have the opportunity during their time in the U.K. to gather with academics and information professionals so that the fellows can explain the ambitions of their research and share discoveries about original material encountered and being cataloged.
- Terms used to describe investigations in the archive are routinely logged and being fed into the taxonomy work.

B.   Accomplishments

The focus of the NEH HCRR grant was planning for the cataloging, catalog enrichment, and transcription work of the GPP through the collaboration of the initial OI and King's fellows' work at the Royal Archives and the teams at the OI, William & Mary Libraries, King's, and the RCT with the following goals.

- Develop taxonomy to support access across a range of descriptive standards, LOC, UKAT*, (UDC)*, specialized vocabulary, and transatlantic language variants.

Following the handover of early research work on automated handwritten transcription from King's to William & Mary, the Foyle Foundation-funded post at King's was able to focus on both the metadata and taxonomy dimensions in the proof-of-concept work and a rapid build of the taxonomy. The NEH HCRR grant enabled this work by supporting the transcription work at William & Mary and the OI, while the work at King's sufficiently impressed the Foyle Foundation to fund the continued work of the Foyle post at King's for a third year.

The GPP partners are working with or have worked with colleagues at the American Antiquarian Society, British Library, Royal Historical Society, the British Institute of Historical Research, the Worshipful Company of Goldsmiths, the Wellcome Institute, and AIM25 (a partnership of 150 institutions holding archives in London) among others in the taxonomy development.

In May 2018, the month after the HCRR grant period concluded, the GPP program manager and academic lead from King's and GPP colleagues from the OI and William & Mary participated in the 2018 *WMQ*-EMSI Workshop—"Archives-based Digital Projects in Early America." (https://oieahc.wm.edu/events/workshops/wmq-emsi/) The workshop produced good contacts for specialist vocabularies that are being followed up.

- Create a metadata production platform that allows participants to crosswalk/harmonize national and international descriptive standards.

The proposed metadata production platform and workflows were envisioned as the key tool for the collaboration amongst the project partners for metadata and transcription. During the grant

period, we encountered some challenges, both technical and organizational, in actualizing the planned workflow and platform.

Security walls for the GPP partners in the U.S. and U.K. meant that the partners engaged in trials and testing over the 12-month grant period to establish a content and metadata delivery workflow.

The establishment of process cycles proved a challenge due to timetables and coordination of the access to content and project groups for digitization, transfer of content, fellows and Royal Archives catalogers.

On September 5, 2017, workshops focused on transcription and platform requirements took place at King's College London with all GPP partners in attendance. After these workshops it emerged that collaboration on metadata development would be reconsidered and go in multiple directions:1) automated fuzzy logic matching against the data; 2) text analysis on transcribed text; and 3) collaboration among scholars and information professionals on metadata might be better placed further in the process as alignment of timetables proved too difficult.

Though the platform requirements workshop session, led by King's Digital Lab (KDL), focused on ideas for the final GPP platform, the event allowed broader discussion on the production versus the final platform capabilities and functions for each. Reaching an understanding on priorities for each platform and technical limits and balancing metadata development (harmonization of standards and vocabulary) and transcription within a single platform raises the difficulty level on the tech and management areas of the system. Priority was given to transcription while challenges to the metadata collaboration were worked on.

Transcription planning and work was aided by consultations with the lead developers of Transkribus, the EU-funded automated handwritten transcription tool. Debbie Cornell of William & Mary Libraries attended the Transkribus conference in Vienna in Fall 2017 and is in regular contact with Transkribus. There was also a Transkribus session for GPP partners in London during the grant period.

Following the September workshop sessions and with continued discussion among the GPP partners, a working group including representatives from the OI, William & Mary, King's, and the Royal Archives is addressing the metadata questions that have arisen during the planning period.

During the grant period, the Roy Rosenzweig Center for History and New Media launched Omeka S, built differently than Omeka (now Omeka Classic) the platform is capable of supporting linked data and multi-site interfaces. The transcription function is yet to be developed for Omeka S, but its metadata capabilities appeared to align with the GPP. As the U.S. partners (William & Mary and OI) focused on transcription, King's Digital Lab (KDL) piloted a short proof-of-concept using Omeka S. Upon evaluation of the Omeka S proof-of-concept, it was determined the multi-level description required for the GPP was not available at the time of testing. The GPP partners came away realizing that significant computer programming would be required to accommodate both functions to enable Omeka or Omeka S to

support the project's metadata goals. Thus, the metadata working group is in the process of finding a solution with KDL taking the lead.

The planning period has now confirmed the use of Omeka for only the GPP transcription platform. The transcription work is ongoing with leadership from William & Mary Libraries in close collaboration with the OI, with the transcribed text being used to further descriptive standards and metadata under the leadership of KDL.

- Test automated tools that speed up transcription.

The testing of automated tools to speed up transcription took place during the NEH HCRR grant period. Transkribus was the automated tool tested, including its Handwritten Text Recognition (HTR) and its OCR capabilities. The OCR capabilities of Transkribus proved fruitful, using the text of manuscript documents that were previously published (*The Correspondence of King George the Third from 1760 to December 1783; Volume 3, July 1773-December 1777* edited by Sir John Fortescue, 1928) allowed us to do Transkribus preparation work in advance of receiving the manuscript images. Prep work involved the cleaning up of OCR text of the documents that will then be used in Transkribus's text2image matching function, once the manuscript images are made available to the OI and William & Mary.

Transkribus HTR was encouraging on the test set of documents, though the initial model building produced a high error rate in consultation with the READ Project. Transkribus identified a few issues with sections of the test set, where low-quality handwriting and ink bleed probably lowered the readability of the images. The transcription team adjusted the parameters for selection of the materials for HTR to eliminate documents with these characteristics and will continue working with Transkribus on collections that meet the new parameters and are not considered appealing for crowd transcription methods.

As a result of the planning period, the key collection deliverable is the transcription of the George III essays that are now available for GPP partners to explore digital methods for scholarship and metadata enrichment. Going forward, a new Transcribe Georgian Papers site will launch in late summer 2018 and enable academics to lead transcription of specific collections within the Georgian Papers in concert with the work of GPP partners and the use of Transkribus.

C.   Audiences

During the NEH HCRR grant period, the GPP reached audiences of scholars and the public in the U.S. and U.K. and around the world through the GPP websites offering access to the digitized materials and the complementary academic interpretation. Websites in the U.S. (https://gpp.wm.edu/) and U.K (http://georgianpapersprogramme.com/) provide academic interpretation and information in concert with the RCT site.

In both the U.S. and the U.K., 26 academic workshops, conference panels, symposia, and public lectures occurred during the grant period, including the major activities outlined below.

In Fall 2017, GPP fellows gathered on September 4, 2017, at Windsor for a symposium to present their research and work in the archives to scholars, librarians, archivist, and technologists working on the GPP. (https://georgianpapers-us.wm.edu/conferences-workshops/workshops/).

In November 2017, the King's Friends was launched as a feature of the GPP designed for researchers (http://georgianpapersprogramme.com/kings-friends/). The King's Friends, which now totals nearly 400 members, is a free-to-join international community of those whose work stands to benefit from the GPP.

On December 1, 2017, a GPP symposium was held at the Library of Congress's Kluge Center to provide scholars and the general public with an opportunity to hear from GPP partners and fellows who have worked in the archives and learn more about the opportunity provided by this rich resource. (Webcast link: https://www.loc.gov/today/cyberlc/feature_wdesc.php?rec=8298)

D.   Evaluation

Representatives from the OI and William & Mary worked in close collaboration and communication with colleagues in the U.K. at the RCT and King's throughout the NEH HCRR grant period to evaluate progress and determine next steps. Work occurred in scheduled GPP meetings and informal communications between meetings.

During the grant period, we encountered challenges, both technical and organizational, in actualizing the planned workflow and platform laid out in the original proposal to the NEH. After evaluating the best options, it was determined that King's Digital Labs would take the lead in developing the plug-ins needed for the metadata production platform and William & Mary Libraries, working with scholars at the OI, will lead on transcription.

One of the challenges of the GPP—and not fully appreciated in the initial phase of the transatlantic partnership—is working cross-culturally with colleagues in the U.S. and U.K. The conventions are different in all ways. This difference in conventions was something that we did not appreciate quite as fully in the early stages of the GPP and have now addressed and accommodated in our planning and work together on both sides of the Atlantic going forward.

E.   Continuation of the Project

The Georgian Papers Programme began as a five-year interdisciplinary project to digitize, conserve, catalog, transcribe, interpret and disseminate an estimated 350,000 items in the Royal Archives and Royal Library relating to the Georgian period, 1714–1837. More detailed work accomplished during the course of the HCRR grant supported a better idea of the number of pages and items involved and the benefits that will accrue from sustained work on metadata enrichment and academic interpretation. On July 17, 2018, the GPP Program Steering Committee, following a careful review of evidence of timing and purposes of the GPP agreed that the GPP should be extended for a further five years and that a revised number of pages, 425,000, be accepted as being in the scope of the project. Memos of understanding between the GPP's primary partners in the U.S. and U.K. will be updated accordingly to recognize the extension of the project's timeframe and the total number of items.

The partnership between RCT and King's and joined by primary United States partners the OI and William & Mary is committed to providing a unique digital resource that is both readily accessible to members of the public and capable of sophisticated manipulation by researchers in any discipline and offers academic and public programming representing fresh research and interpretation.

The GPP draws on the research expertise of hundreds of scholars from a number of disciplines to support the enrichment of catalogues and search tools and to interpret and disseminate the 65,000 items. These materials offer an unparalleled look at the worlds and people who were part of the cultural, economic, political and military realms of the British Empire and beyond. Early exploration has confirmed the Georgian papers' significance for gaining a better understanding of wide-ranging aspects of Atlantic, African, North American, British and European history and producing findings that will have a bearing on our world today.

The GPP is also seeking to build on the work of the OI as a hub for information of relevance to eighteenth-century studies. Presently this is conceived as a three-part spreadsheet or register on the OI and King's websites to include digital projects of interest, research projects and interests that intersect with the GPP and a third category to include other information of interest such as the location of original archives not included in national registers of archives.

F.   Long-Term Impact

In this long-range initiative, the GPP will digitize, disseminate and interpret approximately 425,000 pages or 65,000 items from the reigns of George I through William IV (c. 1714–1837) from the Royal Archives in the Round Tower at Windsor Castle.

Giving scholars the ability to have access to and interpret these newly available documents will enhance and revise the study and interpretation of the American Revolution as we approach 2026 and the 250th anniversary of America's founding. The goal of the project is ultimately to provide a richer understanding and interpretation of the period for the public at large.

As we complete a third full year of work on the GPP, we have articulated the program's pillars as:

1. *Welcoming the World to a Unique Archive*

The GPP is making the archives fully and freely accessible to an international public through conservation, digitization and cataloguing, transcription and powerful search capabilities, and scholarly interpretation. Major exhibitions to support public enjoyment of and access to the archives and teaching materials for schools and universities are in development.

2. *The Past at the Leading Edge*

One of the world's venerable institutions, the British Crown, through the Georgian Papers Programme, is leading in innovative approaches to digitizing, promoting and opening up precious archives from the iconic 12th-century Round Tower of Windsor Castle.

3. *The Global Georgians*

The Georgian papers, 1714–1837, come to us from the threshold of modern history, when science, exploration and trade had created dynamic global connections with transformative consequences. The GPP explores the fully global dimensions of the period.

4. *Creating New Knowledge*

The Georgian papers, a unique and multi-faceted archive, is expanding scholarly exchange across disciplines, generations and among universities, archivists and librarians and researchers beyond the academy, its new discoveries, perspectives and innovative approaches and is creating an important new hub for eighteenth-century studies. Scholarly and public events share these discoveries and new knowledge more broadly.

G.   Award Products

During the grant period, GPP websites in the U.S. and U.K. were updated with blog posts by fellows working in the Royal Archives as well as with information for scholars interested in the project. (https://georgianpapers-us.wm.edu/)   The GPP also presented a number of workshops and public programs, website links of the key programs are included in the appendices.

****

Appendices

Program for Symposium for GPP Fellows, September 4, 2017, Windsor, England:
https://georgianpapers-us.wm.edu/conferences-workshops/workshops/

Transcribed Essays: https://georgianpapers-us.wm.edu/research/transcribed-essays/

Program for GPP Symposium at Library of Congress, December 1, 2017: webcast link at:
https://www.loc.gov/today/cyberlc/feature_wdesc.php?rec=8298

Crowd-source transcription and transcription guidelines (protocols):
http://transcribegeorgianpapers.wm.edu/